Atmospheric Research 211 (2018) 12–20

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

How accurate are the performances of gridded precipitation data products over Northeast China?

Muhammad Abrar Faiz[a], Dong Liu[a,b,e,f,*], Qiang Fu[a,*], Qiaohong Sun[c], Mo Li[a], Faisal Baig[d], Tianxiao Li[a], Song Cui[a]

[a] School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin, Heilongjiang 150030, China
[b] Key Laboratory of Effective Utilization of Agricultural Water Resources of Ministry of Agriculture, Northeast Agricultural University, Harbin, Heilongjiang 150030, China
[c] State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, China
[d] Department of Agricultural Engineering, Bahauddin Zakariya University Multan, Pakistan
[e] Heilongjiang Provincial Collaborative Innovation Center of Grain production Capacity Improvement, Northeast Agricultural University, Harbin, Heilongjiang 150030, China
[f] Key Laboratory of Water-Saving Agriculture of Ordinary University in Heilongjiang Province, Northeast Agricultural University, Harbin, Heilongjiang 150030, China

ARTICLE INFO

Keywords:
Precipitation
Concentration index
Climate change
Gridded datasets

ABSTRACT

The assessment of changes in the precipitation over a region is an important task that is significant for climate modeling and for understanding and monitoring natural hazards and hydrological cycle characteristics. However, the accuracies of gridded precipitation data sets should be evaluated and examined in reference to ground information and observations, as these data can be affected by systematic errors. Therefore, the current study was carried out to evaluate the performances of long-term gridded data sets (from the Global Precipitation Climatology Centre (GPCC), Climate Research Unit (CRU) and University of Delaware (UDEL)) with station-based precipitation data obtained from the National Meteorological Administration (NMA) over Heilongjiang Province in Northeast China for the period from 1961 to 2005. For this purpose, this study adopted spatial comparisons among the data sets, the precipitation concentration index and drought frequency analysis based on the standardized precipitation index to evaluate the performances of these data sets in the entire study area. The results indicate that on spatial scales, the gridded data sets showed higher amounts of annual precipitation at most stations at the spatial scale than the NMA. The spatial correlations among the GPCC, UDEL and NMA were in good agreement compared with the CRU. The annual and seasonal precipitation concentration indexes suggested that the gridded data sets and station-based precipitation exhibited an irregular pattern, as 97% of the stations showed an extremely irregular pattern in the study area during the selected time period. The seasonal results of the precipitation concentration index showed that the gridded data products indicated that all of the stations exhibited irregular precipitation distributions (except during the summer), while the in situ station data indicated that most of the stations had a uniform pattern in the winter and summer. In the summer, the UDEL results were more similar to those of the NMA than those of the CRU and GPCC. The analysis based on gridded data sets revealed that 22–41% of the stations experienced a non-significant negative trend compared with the observed precipitation data sets. The drought frequency analysis based on in situ and gridded data was comparable. This study also revealed a new challenge regarding the utilization of gridded data sets due to discrepancies in the results. Therefore, it is very important to minimize these discrepancies by quantifying their uncertainties, identifying raw data systematic error structures, and applying satellite-based data to improve the quality of data in the region.

1. Introduction

The Intergovernmental Panel on Climate Change reported that climate change has considerable impacts on socioeconomic, physical and biological processes (IPCC, 2014, 2013). These impacts have forced social and scientific communities to understand the consequences and causes of this bitter reality. Precipitation is the most important and active variable in climate dynamics. Changes in atmospheric circulations and precipitation affect the whole ecosystem (Solomon et al., 2007). During the 20th century, the global mean annual precipitation

* Corresponding authors at: School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin, Heilongjiang 150030, China.
E-mail addresses: liudong@neau.edu.cn (D. Liu), fuqiang@neau.edu.cn (Q. Fu).

https://doi.org/10.1016/j.atmosres.2018.05.006
Received 22 January 2018; Received in revised form 2 April 2018; Accepted 3 May 2018
Available online 18 May 2018
0169-8095/ © 2018 Elsevier B.V. All rights reserved.

exhibited a small upward trend, while a larger change was observed in the extreme precipitation than in the mean precipitation (Solomon et al., 2007). Changes in precipitation have a large effect on the water cycle and the safety of the water supply (Faiz et al., 2018, 2017b; Miao et al., 2010; Sun et al., 2016). Therefore, an analysis of precipitation at a monthly resolution is very important because it has a large temporal variability. Due to large temporal irregularities in precipitation, many environmental phenomena, including soil erosion, landslides, floods, drought, and the clogging of reservoirs by sediments, are also linked to these anomalies.

Climate data are very important to understanding and identifying the changes in precipitation characteristics (i.e., the duration, amount, intensity, and type) and variabilities at the regional and global scales (Feng et al., 2004). To investigate climate change, efforts are usually taken to obtain suitable information (Yatagai et al., 2009). Mostly, point data are interpolated over a grid to quantify the variables on different spatial scales. A number of gridded climate data sets are available around the world. For example, the Global Precipitation Climatology Project (GPCP) and the Climate Prediction Center's Merged Analysis of Precipitation (CMAP) data sets both have a resolution of 2.5° × 2.5° (Adler et al., 2003; Xie and Arkin, 1997). According to the CMAP dataset, the mean annual precipitation estimated at the global scale was 1.95 mm × day^{-1} (Gruber and Levizzani, 2008), while it was 2.61 mm × day^{-1} with the GPCP dataset (Adler et al., 2003). This difference in the mean annual precipitation may be due to the coarse resolutions of precipitation data products. Coarse-resolution data sets are less suitable for climate research. Therefore, high-resolution data sets with different statistical approaches have also been developed to obtain useful information for disaster mitigation, initializing numerical models, validating model forecasts and resolving the diurnal cycle of precipitation worldwide (Al et al., 2004; Hijmans et al., 2005). Several gridded data products, including the precipitation reconstruction (PREC), Global Precipitation Climatology Centre (GPCC), CPC Gauge-Based Analysis of Global Daily Precipitation (CPC-Global), Climatic Research Unit (CRU), Climate Prediction Center (CPC), CMAP, GPCP, the University of Delaware (UDEL) products, are available at daily and monthly scales (Chen et al., 2002; Schamm et al., 2014; Xie et al., 2007); these products have a finer resolution of 0.5° × 0.5° and provide useful information such as model forecasts of the hydrological cycle and estimates of climate change (Feng et al., 2011; Marengo et al., 2008; Wen, 2006).

The use of data sets at the daily, monthly, seasonal and annual scales could increase the existing knowledge of the temporal structure and variability, but they cannot provide information regarding gaps in the precipitation, rainy days and the frequencies of time series. Nevertheless, several statistical approaches and indexes are available for precipitation analysis. For instance, the use of the precipitation concentration index (CI) (Martin-Vide, 2004) provides information on the long-term variability and the contribution of high-precipitation days to the total amount. A number of researchers have used the CI to investigate the precipitation concentration worldwide (Benhamrouche et al., 2015; Li et al., 2011; Monjo and Martin-Vide, 2016). For example, Alijani et al. (2008) applied the CI to daily rainfall trends in Iran and found a risk of extreme rainfall in 20% of the country. Zhang et al. (2009) used the CI for the eastern and western river basin of the Zhujiang River for the period from 1960 to 2005 and found an increase in the CI in the western river basin. De Luis et al. (2011) investigated the connection between the precipitation concentration index (PCI) and the annual precipitation trends in Spain and found a distinct behavior that exists between the seasonal PCI and annual precipitation. Coscarelli and Caloiero (2012) evaluated the daily and monthly precipitation concentrations at various spatial and temporal scales in southern Italy and found a clear tendency of the rainfall distribution towards a weaker seasonality. Valli et al. (2013) used the annual and seasonal PCI in Andhra Pradesh, India, for the period from 1981 to 2010 and reported an irregular rainfall distribution. Raziei et al. (2014) used the PCI to study the homogeneity of precipitation and found a regular distribution in the northern parts of Iran. However, such detailed studies for Northeast China are limited. In addition, the drought index is also employed to assess the capability of precipitation products in term of drought frequency analysis. For this purpose, the standardized precipitation index (SPI) (McKee et al., 1993) is used to calculate the drought frequency because calculating the SPI is easy and simple, and it can be used for different time scales.

Heilongjiang Province is situated in Northeast China, and it is one of the fundamental grain-producing bases in the country. Frequent natural disasters (i.e., drought and flooding) have threatened the assets of the region. For example, in the last decade, an average annual economic loss of approximately 227 million USD has been reported due to such disasters in the region (Sun and Bai, 2005). Therefore, it is of great importance to understand the characteristics of such a critical climate variable as precipitation. For this purpose, high-resolution precipitation products should be compared with observed data sets to better understand the characteristics of precipitation. However, these precipitation products are not completely consistent due to distinct generation processes. For example, for the period of 1986–1995, the annual average precipitation amounts obtained from the CRU product and the UDEL product are 697.6 and 697.2 mm × year^{-1}, respectively, while that from the GPCC product is 653.8 mm year^{-1} (Fekete et al., 2004). Therefore, there is an urgent need to understand the differences among gridded precipitation products so that those products could be used to make a comprehensive plan for water management.

The purpose of the current study is to answer certain scientific questions related to gridded precipitation products: (1) How accurate are the performances of different data sets over the study area? (2) How accurate are the capabilities of those data sets at capturing the precipitation pattern and drought frequency? The first objective is to assess the spatial performances of selected gridded precipitation products over the study area. The second objective is to assess the spatial distributions of precipitation data sets based on the precipitation concentration index at the annual and seasonal scales. Finally, a drought frequency analysis is compared with the observed data.

1.1. Research area

The territory of Heilongjiang Province is situated in the northeastern part of China between the longitudes of 120° and 135°E and the latitudes of 45° and 54°N (Fig. 1). The climate of the study area is very cold in the winter and rainy and torrid in the summer. The total annual precipitation ranges from 400 to 700 mm, while the average temperature ranges from 3 to 4.5 °C (Faiz et al., 2017a, 2017b; Khan et al., 2016). The Songhua and Nenjiang Rivers flow through the Sanjiang Plain and Songnen Plain from southwest to northeast. These rivers are a major source of water for drinking, agriculture and industry for the region. Furthermore, the region is rich with grain crops such as sorghum, corn and soybean.

1.2. Data sets and methods

Daily precipitation data from 36 meteorological stations were obtained from the National Metrological Administration (NMA), Heilongjiang Province, China. Three widely used gridded precipitation data products were selected to examine the accuracies of data products in the region. Detailed information on the datasets is presented in Table 1. Historical point data were available from 1961 to 2005; therefore, the gridded data set was also selected for the same period. In this study, the grid cells covering the point data stations were selected to estimate the mean precipitation for each grid cell. The precipitation features were analyzed at a point rather than a spatial context. To obtain the mean precipitation at each grid cell, the MATLAB software was used to convert the latitude and longitude coordinates into grid points. Moreover, ArcGIS 10.2.2 was used for spatial interpolation mapping. In

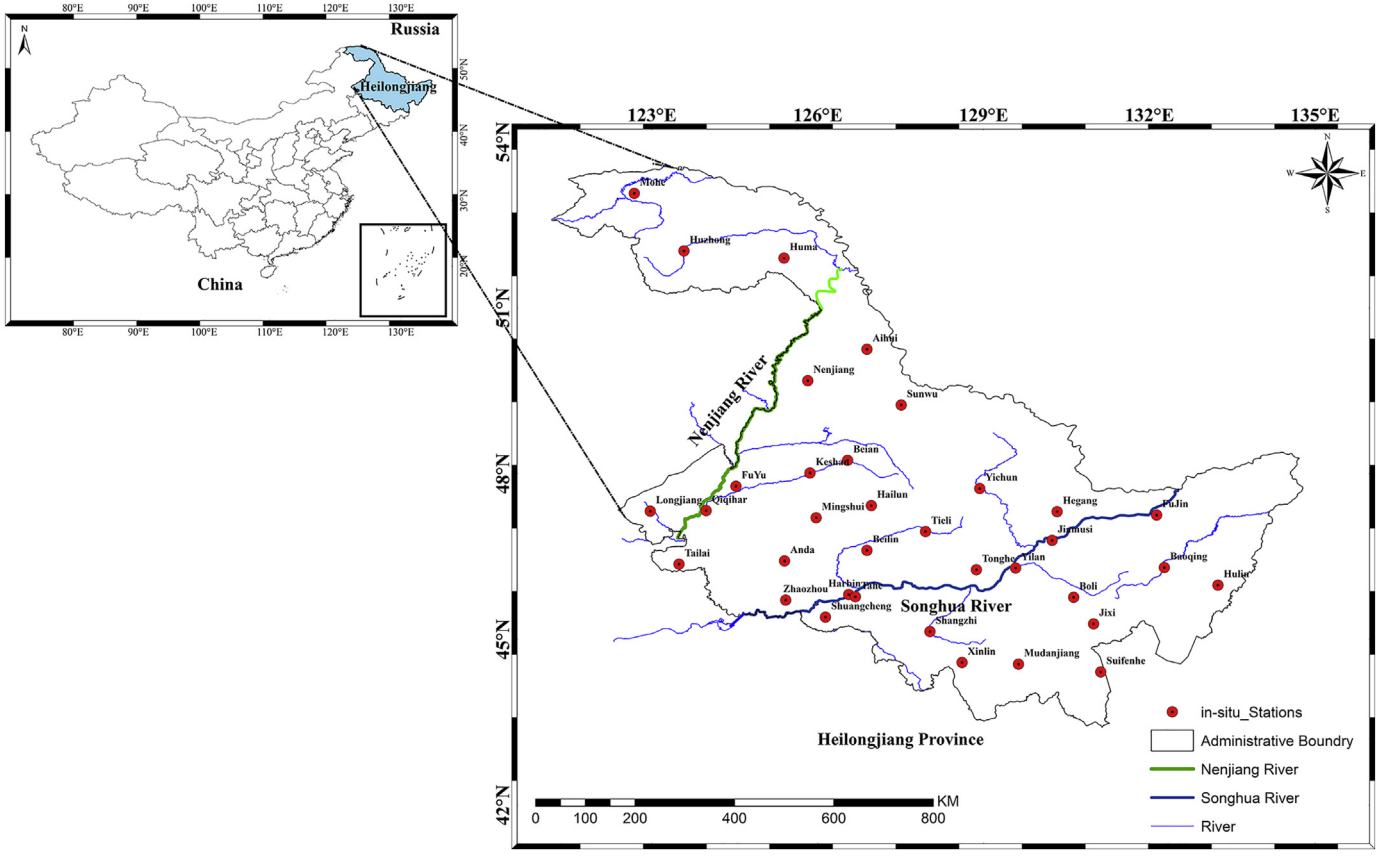

Fig. 1. Location of the study area.

this study, an inverse distance weighted interpolation method was used for the spatial interpolation of the precipitation datasets.

1.3. Precipitation concentration index (PCI)

Based on in situ and gridded precipitation data sets, the precipitation concentration index (PCI) was proposed and developed by (De Luis et al., 1997; Oliver, 1980) and was used at the annual, seasonal and super-seasonal scales. For station and selected grid cell data, the precipitation concentration index (PCI) proposed by (Oliver, 1980) was calculated to find the erosivity and concentration at the annual and seasonal scales. The seasonal and annual precipitation concentrations can be calculated by utilizing the following equations:

$$\mathrm{PCI}_{\mathrm{seasonal}} = \frac{\sum_{i=1}^{3} p_i^2}{\left(\sum_{i=1}^{3} p_i\right)^2} \times 25 \quad (1)$$

$$\mathrm{PCI}_{\mathrm{annual}} = \frac{\sum_{i=1}^{12} p_i^2}{\left(\sum_{i=1}^{12} p_i\right)^2} \times 100 \quad (2)$$

where p_i is the monthly precipitation in month i.

PCI values are more informative for identifying the areas that are prone to flooding or drought compared with annual or seasonal precipitation amounts because PCI values provide insight into the long-term variability in the amount of precipitation received. PCI values of < 10 indicate a low precipitation concentration; a range from 11 to 15 represents a moderate precipitation concentration; values between 16 and 20 represent an irregular distribution of precipitation; and PCI values above 20 display strong irregularities in the precipitation distribution (Royé and Martin-Vide, 2017; Vyshkvarkova et al., 2017; Yeşilırmak and Atatanır, 2016; Zamani et al., 2017a, 2017b).

1.4. Drought frequency analysis

Drought frequency analysis was carried out based on the observed

Table 1
Datasets used in the research.

Dataset	Spatial domain	Temporal domain	Sources	Spatial interpolation
CRU (Climate Research Unit)	0.5° global	Monthly, 1901–2009	CRU, Food and Agriculture Organization, The Global Historical Climatology Network version 2,World Meteorological Organization	The SPHEREMAP method
GPCC (Global Precipitation Climatology Centre)	0.5° global	Monthly, 1901–2010	The Global Historical Climatology Network version 2	Smart interpolation
UDEL (University of Delaware)	0.5° global	Monthly, 1901–2010	GLOBAL SURFACE SUMMARY OF DAY DATA, Global Telecommunication System, The Global Historical Climatology Network version 2, Food and Agriculture Organization, China Meteorological Administration, etc	The SPHEREMAP method

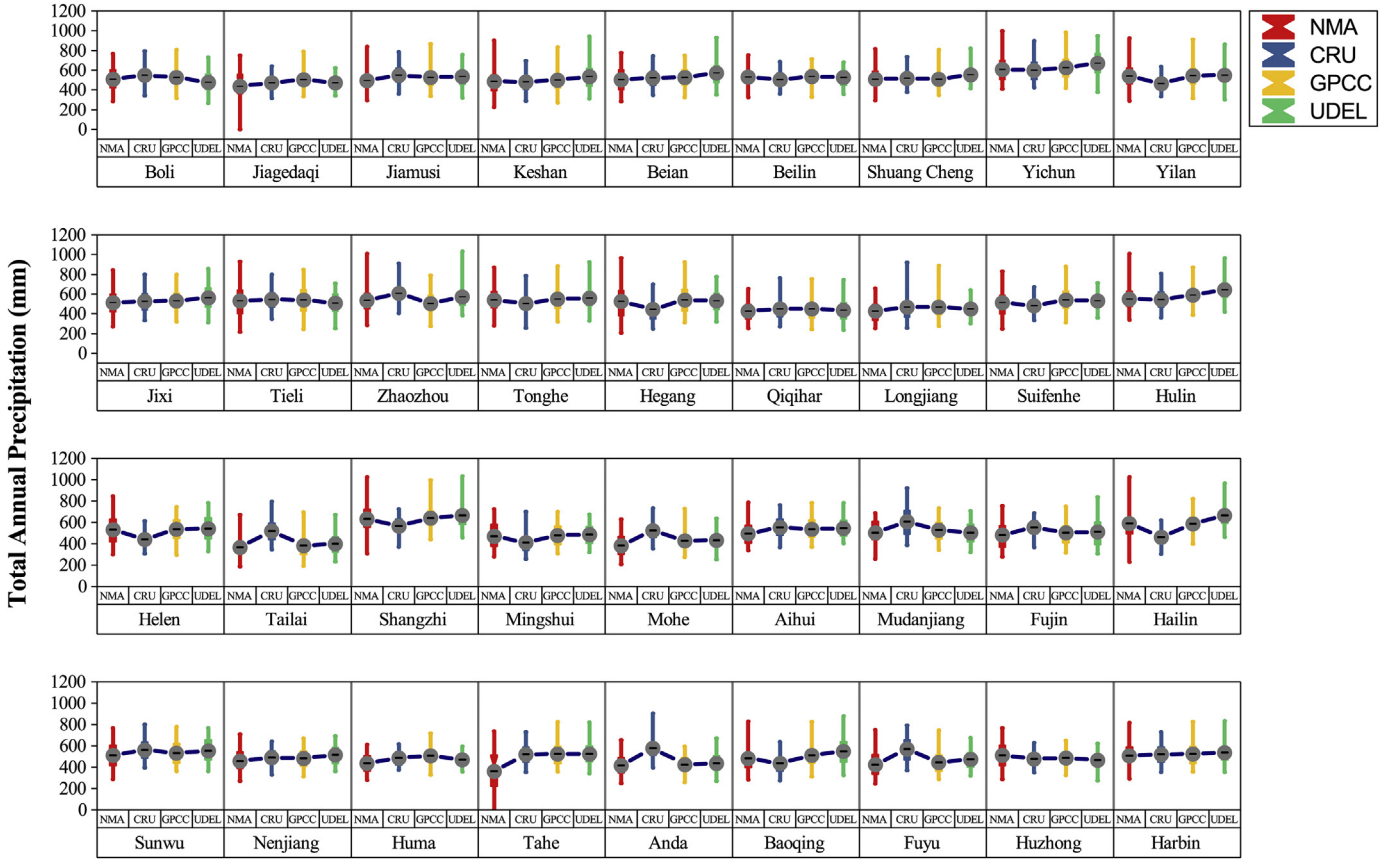

Fig. 2. Box plots for a comparison of precipitation data sets for the period of 1961–2005.

and gridded precipitation in the SPI at the 3-, 6-, and 12-month time scales. The selection of the time scale was made based on the precipitation deficit. For example, hydrological variables such as groundwater, reservoir and stream flow and soil moisture conditions respond to long-term and short-term precipitation anomalies (Svoboda et al., 2012). According to a previous study, the best practical application range of the SPI is from 1 to 24 months (Guttman, 1999). In this study, 3-, 6-, and 12-month time scales were used with SPI ≤ 0 indicating a dry month and SPI ≤ -1.5 indicating a severely dry month. The drought frequency was calculated as the ratio of the number of months with SPI ≤ 0 or SPI ≤ -1.5 to the total number of months. For detailed information regarding the SPI methodology used in this study, readers are referred to the supplementary material.

2. Results and discussion

2.1. Comparison of gridded data products with observations

Fig. 2 presents a comparison of the precipitation data sets from all 36 in situ stations. The results show that the NMA and CRU products had the highest and lowest total annual precipitation at most of the stations, respectively. The UDEL data product indicated an increase in the precipitation during the time period from 1961 to 2005 at 50% of the stations, while the GPCC product showed an increase at 5% of the stations. During the analysis, a small long-term mean difference was observed among the NMA, GPCC and UDEL products compared with CRU over the region. A Taylor diagram (Fig. 3) is also used to examine the pattern of the gridded precipitation with the observed data. A Taylor diagram is a brief summary of an evaluation of the observed and gridded data in terms of the root-mean-square difference, correlation and ratio of variances (Taylor, 2001). Fig. 3 presents the data sets' agreement between the observed and gridded total annual

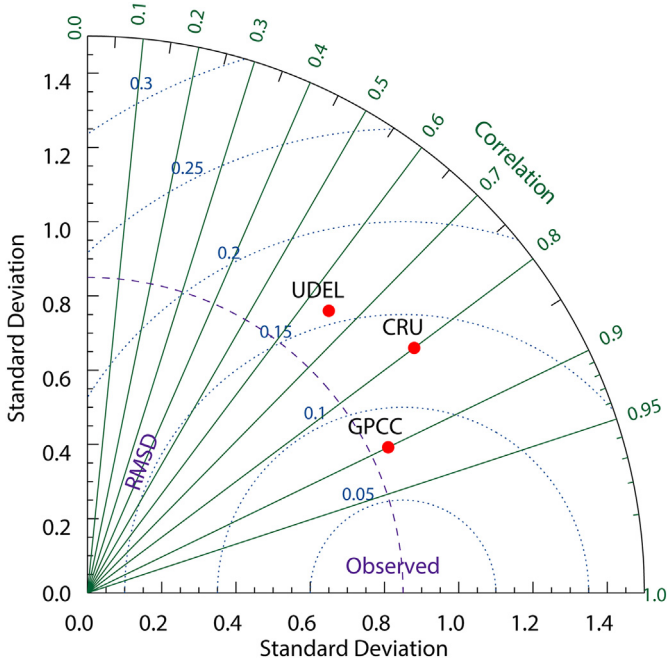

Fig. 3. Taylor diagram for a comparison of the data sets.

precipitation. High values of the root mean square difference (< 0.20) showed good agreement with the annual precipitation dynamics. The correlation coefficient values were concentrated between 0.65 and 0.90. GPCC showed the best agreement with the NMA. In contrast, UDEL was found to be less consistent with the NMA. The higher standard deviation of CRU indicates an intense inter-annual variability

compared with GPCC and UDEL.

Furthermore, to assess the mean rates of change in the regional average precipitation over the study area, a comparison at the temporal scale was also carried out among the precipitation data sets (Fig. S1). The mean rates of change of the regional annual precipitation difference between the NMA and UDEL, CRU, and GPCC products were −3.0, −3.34, and −3.67 mm/decay, respectively. Sequential Mann–Kendall statistics were applied to assess the points of abrupt changes in the precipitation over the study area. The results show that the NMA and UDEL captured the change point in the same year (1995), while the CRU and GPCC precipitation datasets indicated that an abrupt change in the precipitation occurred in the study area in 1994. The gridded datasets had near consistency among themselves. However, gridded data sets depend on ground station observations; in addition, original data are sometimes identical in different databases. Therefore, gridded datasets always show general agreement at various temporal scales. For additional accuracy, more observation gauges should be installed. However, in recent years, a decline in the number of ground base gauge measurements has been seen around the world. For example, in 1901, 10,900 stations from around the world were used for the GPCC full data reanalysis version 7.0 data set. These numbers increased up to 49,470 in 1970, but they reduced to 30,000 in 2005 (Sun et al., 2018). This reduction may be linked to economic and political factors in addition to the migration and abandonment of sites and restrictions on the release of datasets by national meteorological administrations, and this reduction may also be caused by operation costs (Strangeways, 2006). This decline is a serious concern and may represent a big issue in the future when attempting to follow changes in precipitation, and this reduction in the amount of available data may have implications for comparisons with other sources of data.

The spatial distribution of a 45-year long-term average is presented in Fig. 4. Most of the stations based on gridded data sets displayed different characteristics compared with the observed data. The average mean value of the NMA ranged from 183 mm to 644 mm, while those of the CRU, GPCC, and UDEL products ranged from 414 mm to 631 mm, 389 mm to 647 mm, and 403 mm to 670 mm, respectively. Collectively, comparable heavy precipitation events were displayed by all datasets across the Songhua River (Fig. 1). The distribution of long-term differences between the observed and gridded data sets is presented in Fig. 4 (right). The spatial distribution of gridded data sets exhibited lower precipitation in most parts of the region than the NMA, especially across the Songhua River. The analysis also showed that the lower precipitation values in the GPCC, UDEL and CRU products were more significant at most of the stations in the region. This difference in the spatial agreement among the datasets may be due to certain factors. For example, the raw data used to construct the gridded data products were mostly obtained from different sources for each data product (Table 1). In this study, 36 gauge stations were obtained from the NMA, while approximately 200 international exchange stations were used to construct the gridded data products. This indicates that the CRU, UDEL, and GPCC products provide less detailed information over the region (Fig. 4). Furthermore, the gridded data products (CRU, GPCC, and UDEL) also contain climatic data information from adjacent countries to interpolate gridded data. Consequently, adjacent boundaries and raw data may contribute to values that are more negative and relatively low correlation coefficients. In addition to differences in the gridded data products with observational data, quality control and orographic correction factors cannot be ignored. For example, to construct a gridded dataset, raw data are used from different sources that may contain inhomogeneities. Therefore, quality control is necessary to remove unqualified data at the time of implementation of the interpolation process (Ducré-Robitaille et al., 2003). The data sets in this study use different quality control processes. For instance, the first data sets in the GPCC are imported into a rational database, and an elaborate procedure is used to compare the metadata of a station with the metadata of a station that is already available in the database; meanwhile, in the CRU, a two-stage quality control process is adopted before and at the time of interpolation (Rudolf et al., 2011). As seen in Fig. 4 (left), the CRU has a large difference with the NMA compared with the other data sets, which is possibly due to the CRU construction process. A recent study conducted by (Sun et al., 2014) also confirmed a large difference between the CRU and NMA over China.

2.2. Performance evaluation of precipitation products based on precipitation concentration index analysis

The mean values of the precipitation concentration index at an annual scale are presented in Table 2 and Fig. 5. Based on the calculated index values, the range varies among the NMA (18.1 to 26.1), CRU (15 to 24.6), GPCC (15.3 to 24.6), and UDEL (14.8 to 24.9). The precipitation index results suggest that the gridded and station precipitation products exhibited an irregular pattern over all of the stations, and even 97% of the stations showed an extremely irregular pattern in the region during the period from 1961 to 2005. The maximum value of the annual precipitation index for station 23 was 26.6, and the minimum value was 18.1 for station 30, indicating an extremely irregular and uneven dispersion. The calculated values of the index based on the GPCC and UDEL products are consistent with the NMA, while the CRU indicated only an irregular pattern for both stations. Moreover, to make a better conclusion among the data sets, trend analysis of the precipitation data sets was utilized. The results suggest that the GPCC (13) and UDEL (8) stations exhibited a negative precipitation trend compared with the CRU (15) stations (Table 3). The results also depict that, under all of the gridded precipitation data sets, most of the stations had a non-significant trend, and none of the stations exhibited a significant precipitation trend. In contrast, the NMA dataset shows that certain stations (5, 6, 9, 11, 25, 29, and 30) showed a significant increasing trend, while station 13 had a significant decreasing trend in the study area. Overall, the PCI and trend results show that most of the stations exhibited an irregular pattern and a non-significant trend in the region.

The results of the seasonal precipitation concentration index are presented in Tables S1, S2, S3, and S4 and Fig. 6. The seasonal precipitation index value for spring ranged from 7.7 to 16.8 (NMA), 11.2 to 14.9 (UDEL), 11.2 to 15 (GPCC), and 11.7 to 14.6 (CRU) over the region. In spring, approximately all of the stations showed a moderately irregular pattern of precipitation, while only two stations (13 and 29) exhibited a uniform pattern of precipitation in the region. Based on the classification suggested by (Oliver, 1980), the seasonal precipitation index for the summer in situ and gridded data indicate that the NMA (44%), CRU (69%), GPCC (75%), and UDEL (61%) presented uniform distributions of precipitation. During this season, stations near the rivers exhibited patterns with more regularity and precipitation distributions that were more uniform than the other stations (Figs. 1 & 6). Overall, the pattern and dispersion of summer precipitation were more regular than those of spring and fall precipitation. In the winter, the pattern of precipitation was regular, as indicated by in situ station observations compared with the gridded data products. The maximum value of the seasonal precipitation index was observed at station numbers 12 (11.2) and 25 (16.1). In general, the seasonal precipitation had a higher dispersion pattern and a more irregular distribution (i.e., non-uniform) over the area. In winter, the in situ-based stations showed a uniform pattern of precipitation, while the gridded precipitation products indicated an irregular pattern of precipitation over the area. During the analysis, it was observed that seasons with more precipitation showed low values for the seasonal precipitation index, while the low-precipitation seasons, such as the fall and spring, displayed higher precipitation index values. These results were consistent with the findings of other studies (Khalili et al., 2016; Zamani et al., 2017a, 2017b) that examined and investigated the seasonal precipitation concentrations over India and Iran. The results of these studies also suggested that the seasons with the minimum amount of precipitation exhibit a high value of the seasonal concentration index.

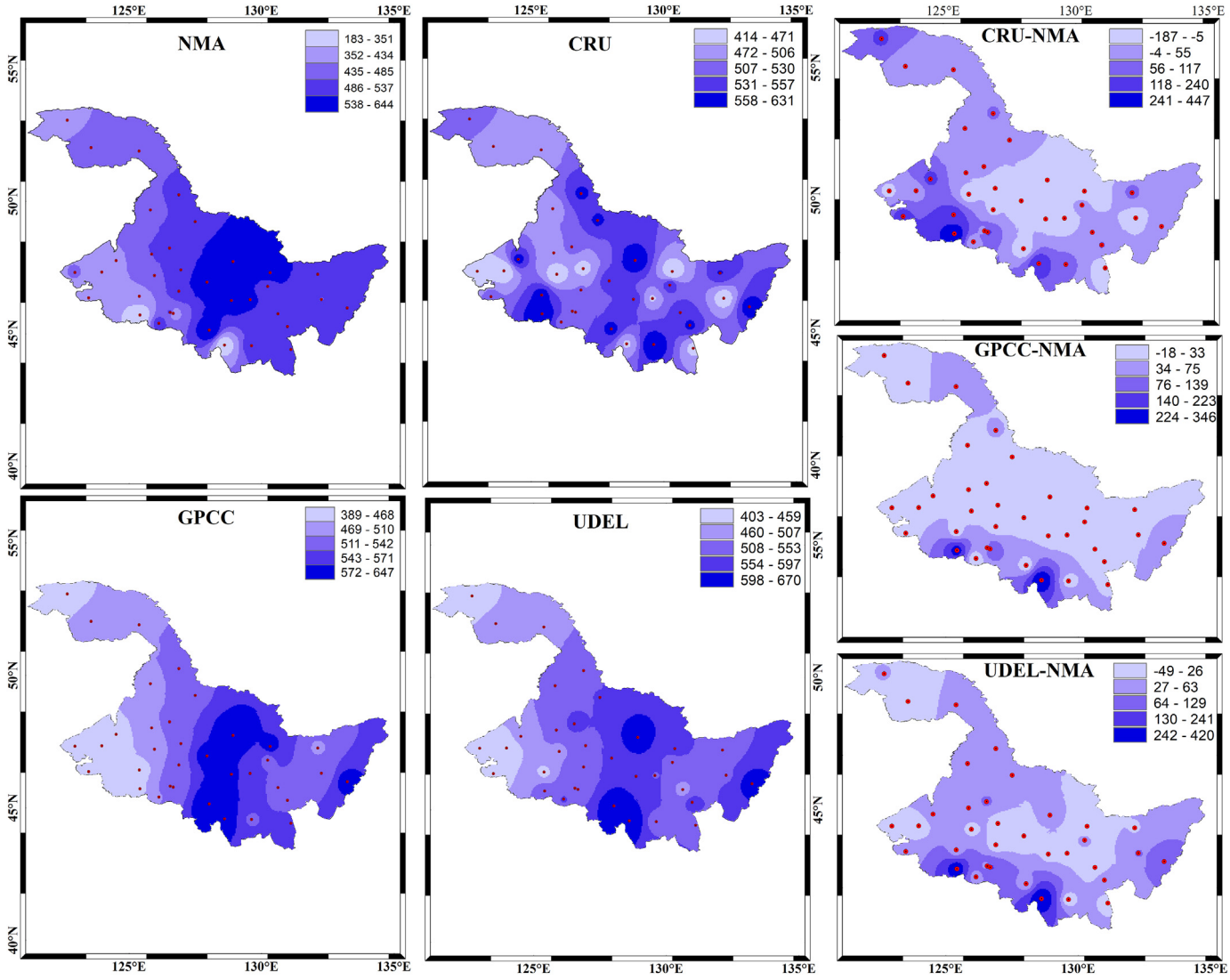

Fig. 4. Spatial distribution of the average precipitation (mm) (left) and the difference (right) between the NMA and the other three gridded precipitation datasets for the period of 1961–2005 in the study area.

Table 2
average values of annual precipitation concentration index at 36 stations for the period of 1961–2005.

ID	Station	NMA	CRU	GPCC	UDEL	ID	Station	NMA	CRU	GPCC	UDEL
1	Yichun	21.4	19.4	19.8	19.1	19	Fujin	20.0	19.3	18.2	19.2
2	Jiamusi	21.5	22.6	19.3	19.5	20	Shangzhi	21.2	22.5	19.1	17.7
3	Yilan	20.9	22.2	19.2	22.1	21	Hailin	25.6	24.3	18.2	21.7
4	Keshan	24.1	20.6	22.0	21.5	22	Mingshui	25.6	21.5	18.2	24.6
5	Jiagedaqi	24.0	18.2	21.3	18.2	23	Tailai	26.6	19.1	23.0	21.0
6	Boli	24.0	20.9	18.3	21.5	24	Helen	22.5	19.5	24.6	18.2
7	Beian	22.6	21.1	20.4	19.3	25	Mohe	21.5	19.0	21.1	19.4
8	Beilin	23.2	21.2	21.3	19.2	26	Aihui	22.0	16.9	18.4	18.3
9	Chuang Cheng	23.2	19.4	20.6	17.3	27	Mudanjiang	20.4	22.3	19.5	17.9
10	Huzhong	23.2	19.3	18.5	19.2	28	Suifenhe	19.9	15.0	18.3	21.7
11	Huma	21.2	21.4	19.7	20.3	29	Zhaozhou	20.7	20.0	16.9	14.8
12	Harbin	23.1	21.4	20.8	20.9	30	Hulin	18.1	20.3	22.5	19.6
13	Tahe	20.4	20.2	20.8	20.2	31	Tonghe	21.4	18.1	15.3	18.7
14	Nenjiang	23.7	19.4	21.1	19.3	32	Tieli	21.8	19.1	19.5	17.8
15	Sunwu	21.7	16.4	18.9	24.3	33	Jixi	21.0	24.6	20.2	18.0
16	Anda	26.0	23.6	24.3	17.9	34	Hegang	22.4	23.5	18.5	23.9
17	Baoqing	19.9	17.6	17.7	23.8	35	Qiqihar	25.5	23.9	20.2	24.9
18	Fuyu	24.7	19.8	22.2	17.9	36	Longjiang	25.5	23.5	24.2	24.6

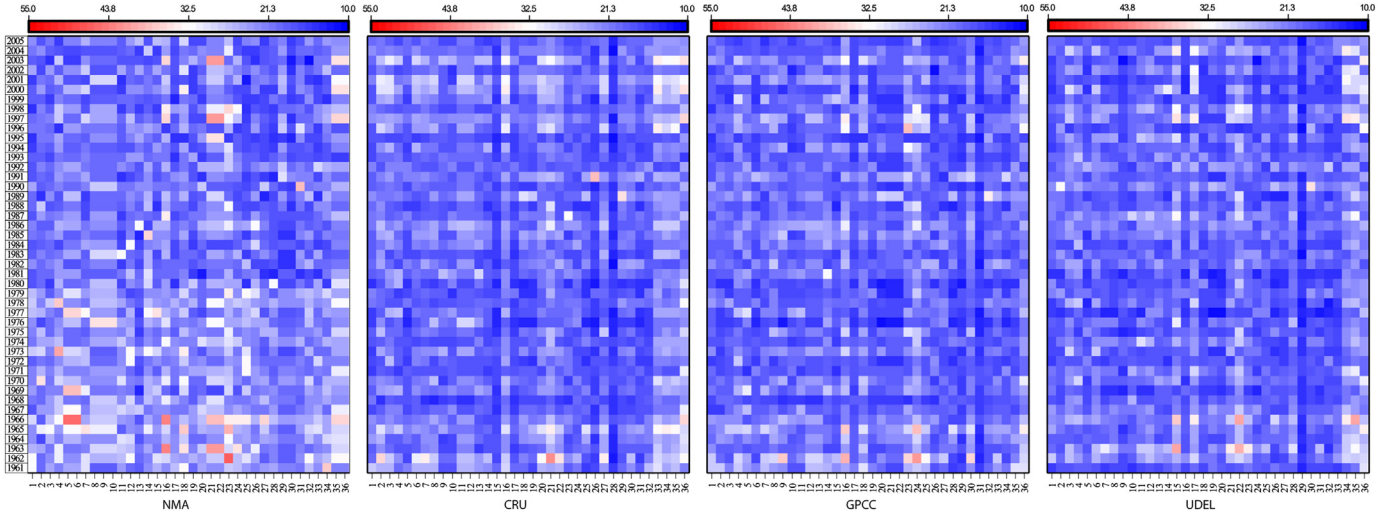

Fig. 5. Precipitation concentration values based on gridded and in situ data sets for the period of 1961–2005.

2.3. Performance evaluation of precipitation products based on drought frequency analysis

The results of the drought frequency analysis based on in situ and gridded data products at 3-, 6-, and 12-month scales are presented in Fig. S2. The results show that all of the stations experienced drought events during the time period of 1961–2005. In this study, the assumption has been made that SPI ≤ 0 indicates a dry month while SPI ≤ −1.5 indicates a severely dry month. The drought frequency is calculated as the ratio of the number of dry or severe (SPI ≤ 0 or SPI ≤ −1.5) months to the total number of months. During the analysis, it was observed at all scales that the drought frequency based on the NMA reached up to 44% for dry months and 6% for extreme months; meanwhile, based on the CRU, GPCC and UDEL, the range of the drought frequency was from 42% for dry months to 6% for extremely dry months. Overall, the SPI based on the gridded and in situ precipitation data showed a good agreement with regard to the drought frequency in the region. In addition, a trend analysis (Hamed and Rao, 1998; Kendall, 1975) of in situ and gridded precipitation at an annual scale was also carried out to assess the trend in the precipitation (Table 3). The results showed that only 5 out of 36 stations based on the NMA experienced a negative trend, while the rest of the stations had a non-significant positive trend in the study area. The trend analysis based on gridded precipitation data products indicates that the CRU revealed 15 out of 36 stations as having a non-significant negative trend, while the GPCC and UDEL presented 13 of 36 stations, and 8 of 36 stations showed a non-significant negative trend (Table 3). Overall, the UDEL had a satisfactory performance with the NMA compared with the GPCC and CRU in capturing the drought frequency and trend analysis.

3. Conclusions

The current analysis was carried out to assess the performances of gridded precipitation data products (i.e., the CRU, GPCC, and UDEL) over Heilongjiang Province, Northeast China. The results demonstrate that the data sets have some differences when compared with in situ data. At different spatial scales, the gridded data sets showed greater amounts of annual precipitation at most of the stations than the NMA. The spatial correlation among the GPCC, UDEL and NMA products showed good agreement compared with the CRU. Compared with the NMA, most of the stations showed higher CRU precipitation differences.

Table 3
Mann Kendall's Z values for each station based on all precipitation products for the period of 1961–2005.

ID	Station	NMA	CRU	GPCC	UDEL	ID	Station	NMA	CRU	GPCC	UDEL
1	Yichun	1.3	-0.22	0.03	-0.22	19	Fujin	0.08	-0.18	-1.05	-0.71
2	Jiamusi	-0.84	-0.84	-0.36	-0.05	20	Shangzhi	1.12	0.32	-0.03	0.03
3	Yilan	1.32	0.13	-0.4	0.32	21	Hailin	-0.32	-0.19	-0.01	-0.13
4	Keshan	1.79	0.22	0.71	1.14	22	Mingshui	0.95	0.67	0.46	0.28
5	Jiagedaqi	4.57	0.46	1.24	1.03	23	Tailai	0.4	-0.66	0.17	0.22
6	Boli	4.65	0.05	-0.54	-0.62	24	Helen	0.5	0.19	0.17	0.4
7	Beian	0.5	-0.42	-0.46	1.05	25	Mohe	3.12	0	1.14	1.3
8	Beilin	0.97	-0.5	0.05	0.03	26	Aihui	1.5	0.62	0.32	1.77
9	Chuang Cheng	2.69	0.22	0.24	0.46	27	Mudanjiang	1.65	-0.24	0.58	0.71
10	Huzhong	1.48	0.19	0.73	0.71	28	Suifenhe	1.79	0.17	-1.12	1.57
11	Huma	1.91	0.11	1.16	0.36	29	Zhaozhou	1.95	-1.26	0.79	0.24
12	Harbin	1.03	0.09	0.19	0.34	30	Hulin	2.24	-0.07	-0.26	-0.32
13	Tahe	-4.02	0.09	0.19	0.01	31	Tonghe	1.12	-0.73	-0.24	0.48
14	Nenjiang	1.3	0.24	0.36	1.42	32	Tieli	-0.3	0.32	-1.09	0.03
15	Sunwu	0.33	-0.44	0.42	0.28	33	Jixi	-0.64	-1.03	-0.5	-0.83
16	Anda	0.64	1.22	0.6	0.13	34	Hegang	1.7	0.3	0.05	0.64
17	Baoqing	0.05	0.03	-1.24	-0.73	35	Qiqihar	0.52	0.85	0.26	0.26
18	Fuyu	0.46	-0.34	0.17	0.24	36	Longjiang	0.67	0.44	0.46	0.21

Highlighted values indication of negative trend.

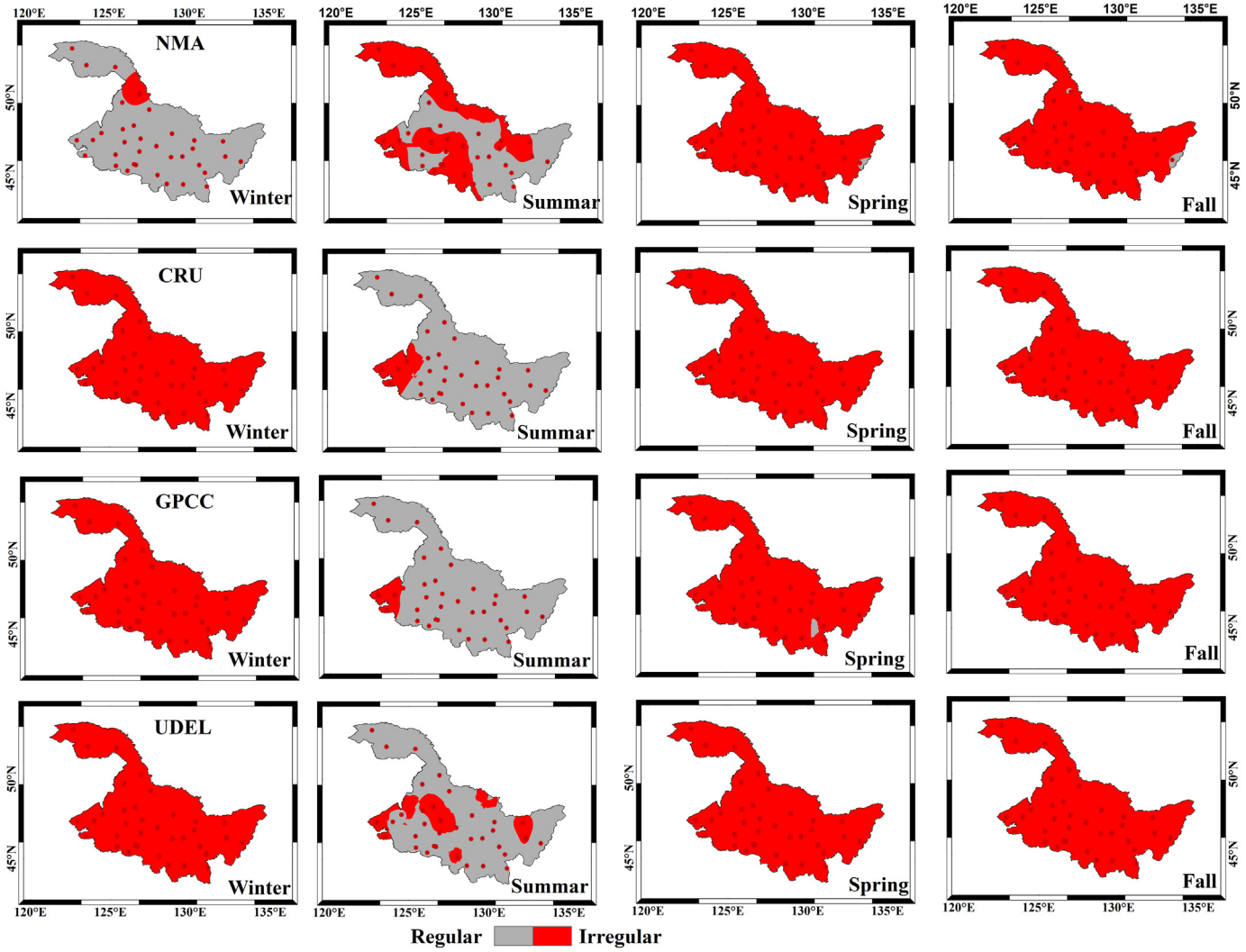

Fig. 6. Spatial distribution maps for the calculated seasonal precipitation index.

The dissimilarities among the data sets may be due to the quality, data source, and techniques used to interpolate the data sets. In addition, the performances of the data sets were also evaluated based on indicators such as annual and seasonal precipitation concentration indexes and drought frequency analysis. The annual precipitation index results suggested that the gridded and station-based precipitation data sets exhibited an irregular pattern over all of the stations, and even 97% of the stations showed an extremely irregular pattern in the region during the period from 1961 to 2005. The seasonal precipitation index was also calculated for gridded and in situ station precipitation. The seasonal results showed that the gridded data products indicated that all of the stations had an irregular precipitation distribution (except during the summer), while the in situ station observations indicated that most of the stations had a uniform pattern in the winter and summer. In the summer season, the UDEL results were more comparable with the NMA results than the CRU and GPCC. The results of the annual precipitation trend indicated that, for the gridded data sets, the CRU displayed 15 out of 36 stations had a non-significant negative trend, while the GPCC and UDEL displayed 13 out of 36 and 8 out of 36 stations, respectively. Meanwhile, based on the NMA, only 5 out of 36 stations experienced a negative trend. The drought frequency analysis based on the in situ and gridded data sets were relatively comparable regardless of whether the spatial and concentration indicator comparisons showed some differences. The results of the current study point out a new challenge regarding the utilization of gridded data sets because a comparison with in situ stations showed some discrepancies. The discrepancies could play an important role in evaluating the performances of hydrological models. These discrepancies may be minimized by quantifying uncertainties, applying raw data systematic error structures, and using satellite-based data to improve the quality of the data.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosres.2018.05.006.

Conflict of interest

No

Acknowledgement

This study is supported by the National Natural Science Foundation of China (No. 51579044, No. 41071053, and No. 51479032), National Key R&D Program of China (No. 2017YFC0406002), Natural Science Foundation of Heilongjiang Province (No. E2017007), Science and Technology Program of Water Conservancy of Heilongjiang Province (No. 201319, No. 201501, No. 201503).

References

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E.,

2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). J. Hydrometeorol. 4, 1147–1167. http://dx.doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.
Al, J.E.T., Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5, 487–503. http://dx.doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
Alijani, B., O'Brien, J., Yarnal, B., 2008. Spatial analysis of precipitation intensity and concentration in Iran. Theor. Appl. Climatol. 94, 107–124. http://dx.doi.org/10.1007/s00704-007-0344-y.
Benhamrouche, A., Boucherf, D., Hamadache, R., Bendahmane, L., Martin-Vide, J., Teixeira Nery, J., 2015. Spatial distribution of the daily precipitation concentration index in Algeria. Nat. Hazards Earth Syst. Sci. 15, 617–625. http://dx.doi.org/10.5194/nhess-15-617-2015.
Chen, M., Xie, P., Janowiak, J.E., Arkin, P.a., 2002. Global land precipitation: a 50-yr monthly analysis based on gauge observations. J. Hydrometeorol. 3, 249–266. http://dx.doi.org/10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.
Coscarelli, R., Caloiero, T., 2012. Analysis of daily and monthly rainfall concentration in southern Italy (Calabria region). J. Hydrol. 416–417, 145–156. http://dx.doi.org/10.1016/j.jhydrol.2011.11.047.
De Luis, M., González-Hidalgo, J.C., Raventós, J., Sánchez, J.R., Cortina, J., 1997. Distribución espacial de la concentración y agresividad de la lluvia en el territorio de la Comunidad Valenciana. Cuaternario y Geomorfol. 11, 33–44.
De Luis, M., González-Hidalgo, J.C., Brunetti, M., Longares, L.A., 2011. Precipitation concentration changes in Spain 1946–2005. Nat. Hazards Earth Syst. Sci. 11, 1259–1265. http://dx.doi.org/10.5194/nhess-11-1259-2011.
Ducré-Robitaille, J.F., Vincent, L.A., Boulet, G., 2003. Comparison of techniques for detection of discontinuities in temperature series. Int. J. Climatol. 23, 1087–1101. http://dx.doi.org/10.1002/joc.924.
Faiz, M.A., Liu, D., Fu, Q., Qamar, M.U., Dong, S., Khan, M.I., Li, T., 2017a. Complexity and trends analysis of hydrometeorological time series for a river streamflow: a case study of Songhua River Basin. China. River Res. Appl. http://dx.doi.org/10.1002/rra.3236.
Faiz, M.A., Liu, D., Fu, Q., Uzair, M., Khan, M.I., Baig, F., Li, T., Cui, S., 2017b. Stream flow variability and drought severity in the Songhua River Basin. Northeast China. Stoch. Environ. Res. Risk Assess. http://dx.doi.org/10.1007/s00477-017-1463-3.
Faiz, M.A., Liu, D., Fu, Q., Wrzesiński, D., Baig, F., Nabi, G., ... Cui, S., 2018. Extreme precipitation and drought monitoring in northeastern China using general circulation models and pan evaporation-based drought indices. Clim. Res. 74 (3), 231–250.
Fekete, B.M., Vörösmarty, C.J., Roads, J.O., Willmott, C.J., 2004. Uncertainties in precipitation and their impacts on runoff estimates. J. Clim. 17, 294–304. http://dx.doi.org/10.1175/1520-0442(2004).
Feng, S., Hu, Q., Qian, W., 2004. Quality control of daily meteorological data in China, 1951–2000: a new dataset. Int. J. Climatol. 24, 853–870. http://dx.doi.org/10.1002/joc.1047.
Feng, L., Zhou, T., Wu, B., Li, T., Luo, J.J., 2011. Projection of future precipitation change over China with a high-resolution global atmospheric model. Adv. Atmos. Sci. 28, 464–476. http://dx.doi.org/10.1007/s00376-010-0016-1.
Gruber, A., Levizzani, V., 2008. Assessment Global Precipitation Report. A Proj. Glob. Energy Water Cycle Exp. Radiat. Panel GEWEX. 16.
Guttman, N.B., 1999. Accepting the standardized precipitation index: a calculation algorithm1. JAWRA J. Am. Water Resour. Assoc. 35, 311–322. http://dx.doi.org/10.1111/j.1752-1688.1999.tb03592.x.
Hamed, K.H., Rao, A.R., 1998. A modified Mann-Kendall trend test for autocorrelated data. J. Hydrol. 204, 182–196.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. http://dx.doi.org/10.1002/joc.1276.
IPCC, 2013. IPCC, 2013. Clim. Chang.
IPCC, 2014. Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press, Cambridge, UK New York, NY.
Kendall, M.G., 1975. Rank. Correl, Meas.
Khalili, K., Tahoudi, M.N., Mirabbasi, R., Ahmadi, F., 2016. Investigation of spatial and temporal variability of precipitation in Iran over the last half century. Stoch. Env. Res. Risk A. 30, 1205–1221. http://dx.doi.org/10.1007/s00477-015-1095-4.
Khan, M.I.M.I., Liu, D., Fu, Q., Dong, S., Liaqat, U.W.U.W., Faiz, M.A.M.A., Hu, Y., Saddique, Q., 2016. Recent climate trends and drought behavioral assessment based on precipitation and temperature data series in the Songhua river basin of China. Water Resour. Manag. 30, 4839–4859. http://dx.doi.org/10.1007/s11269-016-1456-x.
Li, X., Jiang, F., Li, L., Wang, G., 2011. Spatial and temporal variability of precipitation concentration index, concentration degree and concentration period Xinjiang. Chin. Int. J. Climatol. 31, 1679–1693. http://dx.doi.org/10.1002/joc.2181.
Marengo, J.A., Nobre, C.A., Tomasella, J., Oyama, M.D., de Oliveira, G.S., de Oliveira, R., Camargo, H., Alves, L.M., Brown, I.F., 2008. The drought of Amazonia in 2005. J. Clim. 21, 495–516. http://dx.doi.org/10.1175/2007JCLI1600.1.
Martin-Vide, J., 2004. Spatial distribution of a daily precipitation concentration index in peninsular Spain. Int. J. Climatol. 24, 959–971. http://dx.doi.org/10.1002/joc.1030.
McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th Conference on Applied Climatology. American Meteorological Society, Boston, MA, pp. 179–183.
Miao, C., Ni, J., Borthwick, A.G.L., 2010. Recent changes of water discharge and sediment load in the Yellow River basin. China. Prog. Phys. Geogr. 34, 541–561. http://dx.doi.org/10.1177/0309133310369434.
Monjo, R., Martin-Vide, J., 2016. Daily precipitation concentration around the world according to several indices. Int. J. Climatol. 36, 3828–3838. http://dx.doi.org/10.1002/joc.4596.
Oliver, J.E., 1980. Monthly precipitation distribution: a comparative index. Prof. Geogr. 32, 300–309. http://dx.doi.org/10.1111/j.0033-0124.1980.00300.x.
Raziei, T., Daryabari, J., Bordi, I., Pereira, L.S., 2014). Spatial patterns and temporal trends of precipitation in Iran. Theor. Appl. Climatol. 115, 531–540.
Royé, D., Martin-Vide, J., 2017. Concentration of daily precipitation in the contiguous United States. Atmos. Res. 196, 237–247. http://dx.doi.org/10.1016/j.atmosres.2017.06.011.
Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A., Ziese, M., 2011. New GPCC full data reanalysis version 5 provides high-quality gridded monthly precipitation data. Gewex News 21, 4–5.
Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schröder, M., Stender, P., 2014. Global gridded precipitation over land: a description of the new GPCC First Guess Daily product. Earth Syst. Sci. Data 6, 49–60. http://dx.doi.org/10.5194/essd-6-49-2014.
Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., Miller, H.L., 2007. IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. New York Cambridge University Press.
Strangeways, I., 2006. Precipitation: Theory, measurement and distribution. Cambridge University Press, Cambridge.
Sun, Y.G., Bai, R.H., 2005. Research on meteorological disaster in Songhua, Nenjiang River. Meteorol. Press, Beijing (Chinese Tang C, Piechota TC Spat. temporal soil moisture drought Var. Up. Color. River Basin. J Hydrol 379, 122135).
Sun, Q., Miao, C., Duan, Q., Kong, D., Ye, A., Di, Z., Gong, W., 2014. Would the "real" observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. Environ. Res. Lett. 9. http://dx.doi.org/10.1088/1748-9326/9/1/015001.
Sun, S., Chen, H., Ju, W., Wang, G., Sun, G., Huang, J., Ma, H., Gao, C., Hua, W., Yan, G., 2016. On the coupling between precipitation and potential evapotranspiration: contributions to decadal drought anomalies in the Southwest China. Clim. Dyn. 1–19.
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.-L., 2018. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev. Geophys. 56. http://dx.doi.org/10.1002/2017RG000574.
Svoboda, M., Hayes, M., Wood, D., 2012. Standardized Precipitation Index User Guide. World Meteorol. Organ, Geneva, Switz.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192.
Valli, M., Sree, K.S., Krishna, I.V.M., 2013. Analysis of precipitation concentration index and rainfall prediction in various agro-climatic zones of Andhra Pradesh, India. Int. Res. J. Environ. Sci 2, 53–61.
Vyshkvarkova, E., Voskresenskaya, E., Martin-Vide, J., 2017. Spatial distribution of the daily precipitation concentration index in Southern Russia. Atmos. Res. 203, 36–43.
Wen, X.Y., 2006. An overview of China climate change over the 20th century using UK UEA/CRU high resolution grid data. Chin. J. Atmos. Sci. 30, 894–904.
Xie, P., Arkin, P.A., 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull. Am. Meteorol. Soc. 78, 2539–2558. http://dx.doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.
Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., Yang, S., 2007. A gauge-based analysis of daily precipitation over East Asia. J. Hydrometeorol. 8, 607–626. http://dx.doi.org/10.1175/JHM583.1.
Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M.I., Hamada, A., 2009. A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Sola 5, 137–140.
Yeşilırmak, E., Atatanır, L., 2016. Spatiotemporal variability of precipitation concentration in western Turkey. Nat. Hazards 81, 687–704. http://dx.doi.org/10.1007/s11069-015-2102-2.
Zamani, R., Mirabbasi, R., Abdollahi, S., Jhajharia, D., 2017a. Streamflow trend analysis by considering autocorrelation structure, long-term persistence, and Hurst coefficient in a semi-arid region of Iran. Theor. Appl. Climatol. 129, 33–45. http://dx.doi.org/10.1007/s00704-016-1747-4.
Zamani, R., Mirabbasi, R., Nazeri, M., Meshram, S.G., Ahmadi, F., 2017b. Spatio-temporal analysis of daily, seasonal and annual precipitation concentration in Jharkhand state. India. Stoch. Environ. Res. Risk Assess. 1–13. http://dx.doi.org/10.1007/s00477-017-1447-3.
Zhang, Q., Xu, C., Gemmer, M., Chen, Y.D., Liu, C., 2009. Changing properties of precipitation concentration in the Pearl River basin. China. Stoch. Environ. Res. risk Assess. 23, 377–385.